THE ABRAMOV–ROKHLIN ENTROPY ADDITION FORMULA FOR AMENABLE GROUP ACTIONS

Thomas Ward and Qing Zhang

Abstract. In this note we show that the entropy of a skew product action of a countable amenable group satisfies the classical formula of Abramov and Rokhlin.

1. Introduction

Let G be a countable amenable group. We wish to express the entropy of a skew product action of G on a Borel space (defined below) as the sum of a base entropy and a conditional fibre entropy. For G singly–generated, this result was obtained by Abramov and Rokhlin in 1962. Their proof uses two attributes of the acting group: averaging sets (to give convergence in the limit defining conditional fibre entropy) and tiling sets. When the group is singly generated one can choose a sequence of averaging sets that also tile. We describe briefly here what occurs if G is an amenable group. Averaging sets are guaranteed to exist, and the analogous convergence of conditional entropy is obtained by the method that Keiffer used to prove the Shannon–MacMillan theorem for amenable groups in [3]. One cannot (presumably – see [2] and [4] for a description of what is known in this direction) assume the existence of averaging sets that also tile, but the machinery of quasi–tilings developed by Ornstein and Weiss in [4] provides an adequate replacement. The proof below is therefore identical in principle to that of [1], but the arguments to support each step are a little more involved. One specific point should be clarified: we use the deep generalization of Krieger's theorem, due to Rosenthal, which guarantees the existence of a finite generator for a finite entropy free ergodic action of an amenable group. This is not necessary but allows a considerable simplification in the argument. We then show how this implies the general case.

The second author would like to express his appreciation to Professor Benjamin Weiss for his encouragement and advice, and we are both grateful to Professor Dan Rudolph for advice. We also thank an anonymous referee for improvements in the exposition.

1991 *Mathematics Subject Classification.* 28D15, 28D20.

The first author gratefully acknowledges support from NSF grant DMS-91-03056 at the Ohio State University.

2. Quasi–tilings for amenable groups

We now describe the replacement for tiling sets that are needed. The following terminology and results are due to Ornstein and Weiss, [4].
Subsets $A_1, A_2, \dots, A_k$ of G are *ϵ–disjoint* if there are subsets $B_1, B_2, \dots, B_k$ such that

(1) $B_i \subset A_i$ for $i = 1, 2, \dots, k$,
(2) $\frac{|B_i|}{|A_i|} > 1 - \epsilon$, and
(3) $B_i \cap B_j = \emptyset$ for $i \neq j$.

A collection $\{A_1, A_2, \dots, A_k\}$ of subsets of G *α–covers* the set A if

$$\frac{|A \cap (\bigcup_{i=1}^k A_i)|}{|A|} \geq \alpha.$$

A collection $\{A_1, A_2, \dots, A_k\}$ of subsets of G is a *δ–even cover of the set A* if

(1) $A_i \subset A$ for $i = 1, 2, \dots, k$,
(2) there is a number M with $\sum_{i=1}^k \chi_{A_i}(x) \leq M$ for almost every x, and $\sum_{i=1}^k |A_i| \geq (1 - \delta)M$.

Let $K \subset G$ and $\delta > 0$. A subset $A \subset G$ is *(K, δ)–invariant* if

$$\frac{|\{g \in G : Kg \cap A \neq \emptyset \text{ and } Kg \cap (G \backslash A) \neq \emptyset\}|}{|A|} < \delta.$$

Define the *K–boundary of A* to be

$$B(A, K) = \{g \in G : Kg \cap A \neq \emptyset \text{ and } Kg \cap (G \backslash A) \neq \emptyset\}.$$

Lemma 2.1. *If A is (K, δ)–invariant, then for any $c \in G$, the translate Ac is (K, δ)–invariant.*

Proof. It is clear from the definition that $B(A, K)c = B(Ac, K)$, so $|B(A, K)| = |B(Ac, K)|$. □

The property of (K, δ)–invariance is almost preserved under almost disjoint unions in the following sense:

Lemma 2.2. *If the sets A_i, $i = 1, \dots, k$ are (K, δ)–invariant and ϵ–disjoint, then their union $\bigcup_{i=1}^k A_i$ is $(K, (1+\epsilon)\delta))$–invariant. In particular, if the A_i are disjoint, then $\bigcup_{i=1}^k A_i$ is (K, δ)–invariant.*

Proof. It is clear that $B(\bigcup_{i=1}^k A_i, K) \subset \bigcup_{i=1}^k B(A_i, K)$, so

$$|B(\bigcup_{i=1}^k A_i, K)| \leq \sum_{i=1}^k |B(A_i, K)| \leq \delta \sum_{i=1}^k |A_i| \leq \delta(1 + \epsilon)|\bigcup_{i=1}^k A_i|.$$

□

The group G is amenable and therefore admits a Følner sequence, which has the following asymptotic invariance property.

Lemma 2.3. *Let $\{F_n\}$ be a Følner sequence in G. Then, for any finite subset $A \subset G$, and any $\delta > 0$, there is an integer $N > 0$ such that the set F_n is (A, δ)–invariant for all $n \geq N$.*

Proposition 2.4. [4, §1.2] *If $S \subset G$ is a finite set with $e \in S$, and $A \subset G$ is an (SS^{-1}, δ)–invariant set, then the right translates of S that lie in A form a δ–even cover of A.*

Proposition 2.5. [4, §1.2] *If $\{A_\lambda : \lambda \in \Lambda\}$ forms a δ–even cover of A, then there is some $\epsilon > 0$ for which there is an ϵ–disjoint sub–collection of $\{A_\lambda : \lambda \in \Lambda\}$ which $\epsilon(1-\delta)$–covers A.*

For completeness we prove the following theorem (this is proved in [4]).

Theorem 2.6. *Let $e \subset F_1 \subset F_2 \subset \cdots$ be a Følner sequence in G. Then, for any $\epsilon \in (0, \frac{1}{4})$ and any integer $N > 0$, there exist integers $n_1, n_2, \ldots, n_k$ with $N \leq n_1 < n_2 < \cdots < n_k$ such that for any F_M (M sufficiently large), one can find finite subsets $C_1, \ldots, C_k$ of G with the following properties*

(1) *$F_{n_i} C_i \subset F_M$ for $i = 1, 2, \ldots, k$,*
(2) *$F_{n_i} C_i \cap F_{n_j} C_j = \emptyset$ for $i \neq j$,*
(3) *$\{F_{n_i} c : c \in C_i\}$ is an ϵ–disjoint family, and*
(4) *$\{F_{n_i} C_i : i = 1, 2, \ldots, k\}$ forms a $(1-\epsilon)$–cover of F_M.*

Proof. Fix $\frac{1}{4} > \epsilon > 0$ and $N > 0$. Choose $k > 0$ and δ such that $(1 - \frac{\epsilon}{2})^k < \epsilon$ and $6^k \delta < \frac{\epsilon}{2}$. By Lemma 2.3, we can choose $n_1, n_2, \ldots, n_k$ with $N \leq n_1 < n_2 < \cdots < n_k$ such that $F_{n_{i+1}}$ is $(F_{n_i} F_{n_i}^{-1}, \delta)$–invariant and $|F_{n_i}|/|F_{n_{i+1}}| < \delta$. Now for any $(F_{n_k} F_{n_k}^{-1}, \delta)$–invariant F_m with $|F_{n_k}/|F_m| < \delta$, the right translates of F_{n_k} that lie in F_m form a δ–even cover of F_m. By Proposition 2.5, there exists a finite set C_k such that

(1) $\{F_{n_k} c : c \in C_k\}$ is ϵ–disjoint,
(2) $F_{n_k} C_k$ is an $\epsilon(1-\delta)$–cover of F_m, and
(3) $(\epsilon - \delta)|F_m| \leq |F_{n_k} C_k| \leq (\epsilon + \delta)|F_m|$.

To see (3), notice that $|F_{n_k} C_k||F_m|^{-1} > \epsilon(1-\delta) \geq \epsilon - \delta$. On the other hand,

$$|F_{n_k} C_k \backslash F_{n-k} c||F_m|^{-1} \geq |F_{n_k} C_k||F_m|^{-1} - \delta,$$

and $|F_{n_k} C_k \backslash F_{n-k} c||F_m|^{-1} \leq \epsilon(1-\delta)$, so $|F_{n_k} C_k||F_m|^{-1} \leq \epsilon(1-\delta) + \delta \leq \epsilon + \delta$.

Let $D_1 = F_m \backslash F_{n_k} C_k$. we claim that D_1 is $(F_{n_{k-1}} F_{n_{k-1}}^{-1}, 6\delta)$–invariant. Indeed, using Lemma 2.1 and 2.2, we have:

$$\begin{aligned}
|B(D_1, F_{n_{k-1}} F_{n_{k-1}}^{-1})| &\leq |B(F_m, F_{n_{k-1}} F_{n_{k-1}}^{-1})| + |B(F_{n_k} C_k, F_{n_{k-1}} F_{n_{k-1}}^{-1})| \\
&\leq |B(F_m, F_{n_k} F_{n_k}^{-1})| + |C_k||B(F_{n_k}, F_{n_{k-1}} F_{n_{k-1}}^{-1})| \\
&\leq \delta(|F_m| + |C_k||F_{n_k}|) \leq \delta(|F_m| + \frac{1}{1-\epsilon}|C_k F_{n_k}|) \\
&\leq 3\delta|F_m| \leq \frac{3\delta}{1-\epsilon-\delta}|D_1| \leq 6\delta|D_1|
\end{aligned}$$

since $1-\epsilon-\delta > \frac{1}{2}$. It follows that D_1 is $(F_{n_{k-1}}F_{n_{k-1}}^{-1}, 6\delta)$–invariant.

Now consider the size of D_1. It is clear that

$$(1-\epsilon+\delta)|F_m| \geq |D_1| \geq (1-\epsilon-\delta)|F_m|.$$

Since $1-\epsilon > \delta$, $|D_1| > |F_{n_k}| > \frac{1}{\delta}|F_{n_{k-1}}|$, so $|F_{n_{k-1}}||D_1|^{-1} < \delta$. Then there is a finite set C_{k-1} such that

(1) $\{F_{n_{k-1}}c : c \in C_{k-1}\}$ is ϵ–disjoint.
(2) $F_{n_{k-1}}C_{k-1}$ is an $\epsilon(1-6\delta)$–cover of D_1.
(3) $(\epsilon-6\delta)|D_1| \leq |F_{n_{k-1}}C_{k-1}| \leq (\epsilon+6\delta)|D_1|$.

Then let $D_2 = D_1 \backslash F_{n_{k-1}}C_{k-1}$ with

$$|D_2| < (1-\epsilon+\delta)(1-\epsilon+6\delta)|F_m| < (1-\tfrac{\epsilon}{2})^2|F_m|.$$

Inductively, we get D_k with $|D_k| < (1-\epsilon/2)^k|F_m|$ and this implies the theorem.□

¿From now on, we say that sets $A_1,\dots,A_k$ *ϵ–quasi–tile* a set A if there are finite sets $C_1,\dots,C_k$ such that

(1) $A_iC_i \subset A$ for $i = 1,2,\dots,k$,
(2) $A_iC_i \cap A_jC_j = \emptyset$ for $i \neq j$,
(3) $\{A_ic : c \in C_i\}$ forms a ϵ-disjoint family, and
(4) $\{A_iC_i : i = 1,2,\dots,k\}$ forms a $(1-\epsilon)$–cover of A.

The sets $C_1,\dots,C_k$ are called the *tiling centres.*

3. Conditional entropy and entropy

In order to define the entropy of an action of a countable amenable group, an analogue of the total order on the integers adapted to the action is needed; this is furnished by the following Lemma due to Kieffer. The proof is contained in the proof of Lemma 2 in [3]. Notice that the entropy is being implicitly defined as an integral of the information function, and is therefore well–defined without the assumption of ergodicity.

Lemma 3.1. [3] *There is a probability space $(S,\mathcal{S},\lambda)$, a G–action $\{U_g : g \in G\}$ on S and a total order $\prec$ of S such that*

(1) *For each $s \in S$, if $g_1 \neq g_2 \in G$, then $U_{g_1}(s) \neq U_{g_2}$, and*
(2) *for each $g \in G$, $\{s \in S : U_g(s) \prec s\} \in \mathcal{S}$.*

We sketch the proof here for completeness (see [3], page 1033). If G is finite let $S = G$ with uniform measure, and for $\prec$ take any total order on S. Let G act on S by group multiplication. If G is countably infinite, consider the product σ–algebra on $\{0,1\}^G$, and the Bernoulli $(\frac{1}{2},\frac{1}{2})$–measure. Then G acts on $\{0,1\}^G$

by left translation, and we may choose a G–invariant subset $S \subset \{0,1\}^G$ with (1). Restrict the G action to S and order S lexicographically to obtain (2).

For any $s \in S$, one can define a total order $\prec_s$ of G as follows: $g_1 \prec_s g_2$ if and only if $U_{g_1}(s) \prec U_{g_2}(s)$. For any $s \in S$ and $g \in G$, let $V_g(s) = \{g' \in G : g' \prec_s g\}$.

Let $(\Omega, \mathcal{B}, \mu, \{S_g \mid g \in G\})$ be a measure preserving system, so $(\Omega, \mathcal{B}, \mu)$ is a probability space, and $S : g \mapsto S_g$ is an action of G by measure–preserving transformations of $(\Omega, \mathcal{B}, \mu)$.

For any finite measurable partition P of Ω and any subset $A \subset G$, let $P(A)$ denote the smallest σ–algebra containing $S_g^{-1}P$ for all $g \in A$. In particular, $P(\{g\}) = S_g^{-1}P$ for any $g \in G$. For any finite partition P and $\omega \in \Omega$, let $P(\omega)$ denote the unique atom of P containing ω. Now for any sub–σ–algebra $\mathcal{A}$ of $\mathcal{B}$, and any finite partition P, the conditional information function $I(P|\mathcal{A})$ and the conditional entropy $H(P|\mathcal{A})$ can be respectively defined by

$$I(P|\mathcal{A})(\omega) = -\log(\mu)(\{P\}(\omega)|\mathcal{A}).$$

and

$$H(P|\mathcal{A}) = \int I(P|\mathcal{A})(\omega)d\mu.$$

Notice that $H(P|\mathcal{A}) \leq H(P) \leq \log|P|$.

Theorem 3.2. *Let $\{F_n\}$ be a Følner sequence in G with $e \in F_1 \subset F_2 \subset \cdots$ and $F_n \nearrow G$. Then, for any finite partition P and sub–σ–algebra $\mathcal{A}$, the sequence $a_n = \frac{1}{|F_n|} I(P(F_n)|\mathcal{A})$ converges in $L^1(\Omega)$. The limit does not depend on the choice of Følner sequence.*

Proof. From the basic properties of information functions (see [5]), we have

$$\begin{aligned} I(P(F_n)|\mathcal{A})(\omega) &= \sum_{g \in F_n} I(P(\{g\})|P(F_n \cap V_g(s)) \vee \mathcal{A})(\omega) \\ &= \sum_{g \in F_n} I(P|P(F_n g^{-1} \cap V_g(s) g^{-1}) \vee \mathcal{A})(S_g \omega). \end{aligned}$$

Now fix the partition P. For any $E \subset G$, define

$$f_E(\omega, s) = I(P|P(E \cap V_e(s)) \vee \mathcal{A})(\omega).$$

One can check that f_E is a measurable function on $\Omega \times S$. Then

$$I(P(F_n)|\mathcal{A})(\omega) = \sum_{g \in F_n} I(P|P(F_n g^{-1} \cap V_e(U_g s)) \vee \mathcal{A})(S_g \omega) = \sum_{g \in G} f_{F_n g^{-1}}(S_g \omega, U_g s),$$

where U is the G–action given by Lemma 3.1. It is clear that for any sequence $E_1 \subset E_2 \subset \cdots$, $E_n \nearrow G$ the limit $\lim_{n\to\infty} f_{E_n} = f_G$ exists in $L^1(\Omega \times S)$. For any

$\epsilon > 0$, there is a finite set B such that if $E \supset B$, $\|f_E - f_G\|_1 < \epsilon$. Since B is a finite set, when n is sufficiently large we have

$$\frac{|F_n \cap \left(\bigcap_{b\in B} b^{-1}F_n\right)|}{|F_n|} \geq (1-\epsilon).$$

It is clear that for any $g \in F_n \cap \left(\bigcap_{b\in B} b^{-1}F_n\right)$, we have $F_n g^{-1} \supset B$. Therefore, when n is sufficiently large we have

$$\begin{aligned}
\|\frac{1}{|F_n|}I(P(F_n)|\mathcal{A}) - f_G\|_{L^1(\Omega\times S)} &\leq \frac{1}{|F_n|}\sum_{g\in F_n}\|f_{F_n g^{-1}}(S_g\times U_g) - f_G\| \\
&\leq \frac{1}{|F_n|}\sum_{g\in F_n\cap(\cap_{b\in B}b^{-1}F_n)}\|f_{F_n g^{-1}}(S_g\times U_g) - f_G\| \\
&\qquad + \frac{|F_n\backslash F_n \cap \left(\bigcap_{b\in B} b^{-1}F_n\right)|}{|F_n|}\log|P| \\
&\leq \epsilon(1+\log|P|).
\end{aligned}$$

This implies that

$$\lim_{n\to\infty}\|\frac{1}{|F_n|}I(P(F_n)|\mathcal{A}) - f_G\|_{L^1(\Omega\times S)} = 0.$$

Therefore

$$\lim_{n\to\infty}\|\frac{1}{|F_n|}I(P(F_n)|\mathcal{A}) - \int f_G d\lambda\|_{L^1(\Omega)} = 0,$$

and the theorem follows. □

Corollary 3.3. *For any Følner sequence $\{F_n\}$ satisfying $F_1 \subset F_2 \subset \cdots$, $F_n \nearrow G$, the limit*

$$\lim_{n\to\infty}\frac{1}{|F_n|}H(\bigvee_{g\in F_n} S_g P|\mathcal{A})$$

exists and is independent of the choice of $\{F_n\}$.

We will use $h(S,P|\mathcal{A})$ to denote the limit $\lim_{n\to\infty}\frac{1}{|F_n|}H(\bigvee_{g\in F_n} S_g P|\mathcal{A})$ and define the *conditional entropy of S with respect to $\mathcal{A}$* by $h(S|\mathcal{A}) = \sup_P h(S,P|\mathcal{A})$. The *entropy of the G–action S* is defined to be the conditional entropy of S with respect to the trivial sub–σ–algebra $\mathcal{N} = \{\emptyset, \Omega\}$: $h(S) = h(S|\mathcal{N})$. Similarly, we define $h(S,P)$ to be $h(S,P|\mathcal{N})$.

If $\mathcal{A}$ is a finite σ–algebra, let $P(\mathcal{A})$ be the finite partition that generates $\mathcal{A}$.

Lemma 3.4. *If $\{\mathcal{A}_n\}$ is a sequence of finite σ–algebras with $\mathcal{A}_n \nearrow \mathcal{B}$, then $h(S) = \lim_{n\to\infty} h(S,P(\mathcal{A}_n))$.*

Proof. For finite partitions P and Q,

$$\begin{aligned}
H\left(\bigvee_{g\in F_n} S_gP\right) &\leq H\left(\bigvee_{g\in F_n} S_gP \vee \bigvee_{g\in F_n} S_gQ\right)\\
&\leq H\left(\bigvee_{g\in F_n} S_gQ\right) + H\left(\bigvee_{g\in F_n} S_gP | \bigvee_{g\in F_n} S_gQ\right)\\
&\leq H\left(\bigvee_{g\in F_n} S_gQ\right) + \sum_{g\in F_n} H\left(S_gP | \bigvee_{g\in F_n} S_gQ\right)\\
&\leq H\left(\bigvee_{g\in F_n} S_gQ\right) + |F_n|H(P|Q),
\end{aligned}$$

so $h(S,P) \leq h(S,Q) + H(P|Q)$.

An easy consequence of the Increasing Martingale theorem shows that if P is a finite partition, then $H(P|\mathcal{A}_n) \searrow H(P|\mathcal{B}) = 0$ (see [6], page 38). Hence $h(S,P) \leq h(S,P(\mathcal{A}_n) + H(P|P(\mathcal{A}_n))$ and $H(P|P(\mathcal{A}_n)) \to 0$ as $n \to \infty$. It follows that $h(S,P) \leq \lim_{n\to\infty} h(S,P(\mathcal{A}_n))$ for any finite partition P, so $h(S) \leq \lim_{n\to\infty} h(S,P(\mathcal{A}_n))$; the reverse inequality is clear. □

4. Entropy addition formula

Let $(X,\mathcal{B},\mu,\{T_g : g \in G\})$ be a measure preserving system and let $(Y,\mathcal{C},\nu)$ be a probability space. Let $MPT(Y)$ denote the group of all invertible measure preserving transformations of Y and let $\alpha : X \times G \to MPT(Y)$ be a cocycle with the property that for any fixed $g \in G$, $\alpha(x,g)(y)$ is a measurable Y–valued function of x and y with respect to the product σ–algebra $\mathcal{B}\otimes\mathcal{C}$. Let $\Omega = X \times Y$. Define a measure preserving G–action $\{S_g : g \in G\}$ on Ω by:

$$S_g(x,y) = (T_gx, \alpha(x,g)y).$$

The action S is then a *skew–product* extension of T by α. For a set $B \in \mathcal{B}$ ($C \in \mathcal{C}$), we also use B (resp. C) to denote the set $B \times Y$ (resp. $X \times C$) in $\mathcal{B}\otimes\mathcal{C}$. This notational device amounts to a canonical embedding, $\mathcal{B} \hookrightarrow \mathcal{B}\otimes\mathcal{C}$ (resp. $\mathcal{C} \hookrightarrow \mathcal{B}\otimes\mathcal{C}$)

In order to prove the entropy addition formula without the assumption of freeness (see Theorem 4.4 below) we will need an independent proof of the formula for the entropy of a direct product. This may be obtained for group actions exactly as for single transformations (see [6], page 61); we include a short proof for completeness.

If the cocycle $\alpha(x,g)$ is independent of $x \in X$ then $\alpha(x,g) = V_g$ for some G–action V on $(Y,\mathcal{C},\nu)$, and the skew product S above is then the direct product $S_g = T_g \times V_g$.

Lemma 4.1. *The entropy of a direct product is the sum of the entropies:*

$$h(T \times V) = h(T) + h(V).$$

Proof. Let $\{\mathcal{B}_n\}$ and $\{\mathcal{C}_n\}$ be sequences of finite σ–algebras with $\mathcal{B}_n \nearrow \mathcal{B}$ (that is, $\mathcal{B}_n \subset \mathcal{B}_{n+1}$ for all n, and $\bigcup_n \mathcal{B}_n$ generates $\mathcal{B}$) and $\mathcal{C}_n \nearrow \mathcal{C}$. Then, by independence,

$$h(S, P(\mathcal{B}_n \times \mathcal{C}_n)) = h(T, P(\mathcal{B}_n)) + h(V, P(\mathcal{C}_n)).$$

Applying Lemma 3.4 gives the result. $\square$

Theorem 4.2. [7] *If T is an ergodic free G–action with $h(T) < \infty$, then there is a finite partition ξ such that $\mathcal{B} = \bigvee_{g \in G} T_g \xi$.*

Such a partition ξ will be called a *generator* of $(X, \mathcal{B}, \mu, \{T_g : g \in G\})$. In fact Rosenthal proves a much stronger result, exhibiting a finite uniform generator of optimal (least) cardinality.

Proposition 4.3. *Let S and T be the measure preserving G–actions defined above, and assume that the base action T is ergodic and free. Then $h(S) = h(T) + h(S|\mathcal{B})$.*

Proof. We first show that $h(S) \geq h(T) + h(S|\mathcal{B})$. It is enough to show that the supremum of $h(T, P)$ over all partitions P of $X \times Y$ which are of the form $P = \xi \times \eta$, where ξ and η are finite partitions of X and of Y respectively, is greater than or equal to $h(T) + h(S|\mathcal{B})$. Let $\{F_n\}$ be a Følner sequence in G such that $e \in F_1 \subset F_2 \subset \cdots$ and $F_n \nearrow G$. For a partition $P = \xi \times \eta$,

$$H(P(F_n)) = H(\xi(F_n) \vee \eta(F_n)) = H(\xi(F_n)) + H(\eta(F_n)|\xi(F_n))$$

and so

$$H(P(F_n)) \geq H(\xi(F_n)) + H(\eta(F_n)|\mathcal{B}).$$

By Corollary 3.3, we have

$$\begin{aligned} h(S, P) &= \lim_{n\to\infty} \frac{1}{|F_n|} H(P(F_n)) \\ &\geq \lim_{n\to\infty} \frac{1}{|F_n|} H(\xi(F_n)) + \lim_{n\to\infty} \frac{1}{|F_n|} H(\eta(F_n)|\mathcal{B}) \\ &= h(T, \xi) + h(S, \eta|\mathcal{B}) \end{aligned}$$

Now we show that $h(S) \leq h(T) + h(S|\mathcal{B})$. We need only consider the case $h(T) < \infty$. By Theorem 4.2, there is a finite generator ξ for $(X, \mathcal{B}, \mu, \{T_g : g \in G\})$. Let P be any finite partition of $\Omega = X \times Y$. For any $\epsilon > 0$,there is an N such that when $n > N$,

$$|\frac{1}{|F_n|} H(P(F_n)) - h(S, P)| < \epsilon,$$

$$|\frac{1}{|F_n|}H(\xi(F_n)) - h(T,\xi)| < \epsilon,$$

and

$$|\frac{1}{|F_n|}H(P(F_n)|\mathcal{B}) - h(S,P|\mathcal{B})| < \epsilon.$$

By Theorem 1.6, for $\epsilon > 0$ and an integer $N > 0$, there exist $n_1, \dots, n_k$ with $N < n_1 < \cdots < n_k$ for which the sets $F_{n_1}, \dots, F_{n_k}$ ϵ–quasi–tile any F_m with m sufficiently large.

Since ξ is a generator, the Increasing Martingale theorem (see [6], page 38) shows that for any finite partition Q, $H(Q|\xi(F_k)) \searrow H(Q|\mathcal{B})$ as $k \to \infty$. It follows that there is a finite set B such that for any set $A \supset B$,

$$H(P(F_{n_i})|\xi(A)) \leq H(P(F_{n_i})|\mathcal{B}) + \epsilon$$

for $i = 1, \dots, k$.

Now for m sufficiently large, F_m is (B,ϵ)–invariant and can be ϵ–quasi–tiled by $F_{n_1}, \dots, F_{n_k}$.

Now

$$\frac{1}{|F_m|}H(P(F_m)) \leq \frac{1}{|F_m|}H(\xi(F_m)) + \frac{1}{|F_m|}H(P(F_m)|\xi(F_m)).$$

Therefore

$$h(S,P) \leq h(T) + \frac{1}{|F_m|}H(P(F_m)|\xi(F_m)) + 2\epsilon.$$

Let $C_1, \dots, C_k$ be tiling centres for F_m. Then

$$|F_m| \geq |\bigcup_{i=1}^{k} F_{n_i}C_i| \geq (1-\epsilon)|F_m| \text{ and } |\bigcup_{i=1}^{k} F_{n_i}C_i| \geq (1-\epsilon)\sum_{i=1}^{k}|C_i||F_{n_i}|.$$

Now we have

$$\begin{aligned}
\frac{1}{|F_m|}H(P(F_m)|\xi(F_m)) &\leq \frac{1}{|F_m|}H(P(\bigcup_{i=1}^{k} F_{n_i}C_i)|\xi(F_m)) + \epsilon\log|P| \\
&\leq \frac{1}{|\bigcup_{i=1}^{k} F_{n_i}C_i|}H(P(\bigcup_{i=1}^{k} F_{n_i}C_i)|\xi(F_m)) + \epsilon\log|P| \\
&\leq \frac{(1-\epsilon)^{-1}}{\sum_{j=1}^{k}|C_j||F_{n_j}|}\sum_{i=1}^{k} H(P(F_{n_i}C_i)|\xi(F_m)) + \epsilon\log|P|.
\end{aligned}$$

Let $t_i = |C_i||F_{n_i}|/(\sum_{i=1}^{k}|C_i||F_{n_i}|)$ for $i = 1, 2, \dots, k$. Then $1 \geq t_i > 0$, $\sum t_i = 1$, and so

$$\frac{1}{\sum_{j=1}^{k}|C_j||F_{n_j}|}\sum_{i=1}^{k} H(P(F_{n_i}C_i)|\xi(F_m)) = \sum_{i=1}^{k}\frac{t_i}{|C_i||F_{n_i}|}H(P(F_{n_i}C_i)|\xi(F_m)).$$

Since $e \in F_{n_i}$, $C_i \subset F_m$ and $C_i \cap C_j = \emptyset$ for $i \neq j$. Let $A = \{g \in F_m : Bg \subset F_m\}$. Since F_m is (B, ϵ)–invariant, $|A|/|F_m| \geq (1-\epsilon)$. Therefore, for any $1 \leq i \leq k$,

$$\begin{aligned}
\frac{1}{|C_i||F_{n_i}|} H(P(F_{n_i}C_i)|\xi(F_m)) &\leq \frac{1}{|C_i|}\Big(\sum_{c\in C_i} \frac{1}{|F_{n_i}|} H(P(F_{n_i}c|\xi(F_m)))\Big) \\
&\leq \frac{1}{|C_i|}\Big(\sum_{c\in C_i\cap A} \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\xi(F_m c^{-1})) \\
&\qquad + \sum_{c\in F_m\setminus A} \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\xi(F_m c^{-1}))\Big) \\
&\leq \frac{1}{|C_i|}\Big(\sum_{c\in C_i\cap A} \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\xi(F_m c^{-1}))\Big) \\
&\qquad + \frac{|F_m\setminus A|}{|C_i||F_{n_i}|} \log|P| \\
&\leq \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\xi(B)) + \frac{|F_m\setminus A|}{|C_i||F_{n_i)}|} \log|P| \\
&\leq \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\mathcal{B}) + \frac{|F_m\setminus A|}{|C_i||F_{n_i}|} \log|P| + \epsilon.
\end{aligned}$$

This implies that

$$\begin{aligned}
\frac{1}{|F_m|} H(P(F_m)|\xi(F_m)) &\leq \frac{1}{1-\epsilon}\Big(\sum_{i=1}^{k} t_i \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\mathcal{B}) \\
&\qquad + \sum_{i=1}^{k} t_i \frac{|F_m\setminus A|}{|C_i||F_{n_i}|} \log|P|\Big) + \epsilon(1+\log|P|) \\
&\leq \frac{1}{1-\epsilon} \sum_{i=1}^{k} t_i \frac{1}{|F_{n_i}|} H(P(F_{n_i})|\mathcal{B}) \\
&\qquad + \frac{1}{(1-\epsilon)^2} \frac{|F_m\setminus A|}{|F_m|} \log|P| + \epsilon(1+\log|P|)
\end{aligned}$$

Since $n_i > N$, we have

$$\frac{1}{|F_m|} H(P(F_m)|\xi(F_m)) \leq \frac{1}{1-\epsilon} h(S, P|\mathcal{B}) + \epsilon\Big(\big(\frac{1}{(1-\epsilon)^2} + \frac{1}{1-\epsilon} + 1\big) \log|P| + 1\Big).$$

If $0 < \epsilon < \frac{1}{2}$ and $|P| \geq 2$, then

$$\frac{1}{|F_m|} H(P(F_m)|\xi(F_m)) \leq \frac{1}{1-\epsilon} h(S, P|\mathcal{B}) + 8\epsilon \log|P|.$$

Therefore

$$h(S, P) \leq h(T) + h(S, P|\mathcal{B}) + 10\epsilon \log|P|.$$

Since ϵ was arbitary, we have $h(S, P) \leq h(T) + h(S, P|\mathcal{B})$. The theorem follows.$\square$

Theorem 4.4. *Let S and T be the measure preserving G–actions defined above. Then $h(S) = h(T) + h(S|\mathcal{B})$.*

Proof. There are two reductions to be carried out. First, let $T = \int_0^1 T^{(s)} ds$ be the ergodic decomposition for T (this is constructed for any countable group action in [8, §4]). We then have $h(T) = \int_0^1 h(T^{(s)})ds$ (this follows easily from the definition of entropy for G–actions given in Section 3 above). Writing $S^{(s)}(x, y) = (T_g^{(s)}(x), \alpha(x, g)(y))$, we obtain

$$h(S) = \int_0^1 (h(T^{(s)}) + h(S^{(s)}|\mathcal{B}))ds = h(T) + h(S|\mathcal{B})$$

by Proposition 4.3.

We may therefore assume that T is an ergodic action. Define an action U of G as follows. Let $Z = \{0, 1\}^G$ with the Bernoulli $(\frac{1}{2}, \frac{1}{2})$–measure η defined on the σ–algebra of Borel sets $\mathcal{D}$ obtained from the discrete topology on $\{0, 1\}$. The group G acts via the shift, $U_g(\mathbf{z})_h = z_{gh}$ where $\mathbf{z} = (z_g)_{g \in G} \in Z$. An easy calculation shows that $h(U) = \log 2$; moreover U acts freely. To see this, notice that $\{\mathbf{z} \mid U_g \mathbf{z} = \mathbf{z}\}$ has zero measure if either $\{g^n\}$ or $G/\{g^n\}$ is infinite, and one of these must occur unless G is finite – in which case all the entropies are zero. Let G act on $X \times Z \times Y$ via $S_g^U(x, \mathbf{z}, y) = ((T_g \times U_g)(x, \mathbf{z}), \alpha(x, g)(y))$. Then it is clear that $h(S^U|\mathcal{B}) = h(S|\mathcal{B})$ since α is independent of the Z coordinate. Also, the base action $T \times U$ is free, so we may apply Proposition 4.3 and Lemma 3.4 to obtain $h(S) + h(U) = h(S^U) = h(T \times U) + h(S^U|\mathcal{B}) = h(T) + h(U) + h(S|\mathcal{B})$, which gives the result since $h(U)$ is finite. $\square$

References

[1] L. M. Abramov & V. A. Rokhlin, *The entropy of a skew product of measure–preserving transformations*, Amer. Math. Soc. Transl. (Ser. 2) **48** (1965), 225–265.

[2] C. Chou, *Elementary amenable groups*, Illinois Journal of Math. **24** (1980), 396–407.

[3] J. C. Kieffer, *A generalized Shannon–McMillan theorem for the action of an amenable group on a probability space*, Annals of Prob. **3** (1975), 1031–1037.

[4] D. S. Ornstein & B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, Journal d'Analyse Mathématique **48** (1987), 1–141.

[5] W. Parry, *Entropy and Generators in Ergodic Theory*, Benjamin, New York, 1969.

[6] ———, *Topics in Ergodic Theory*, Cambridge University Press, Cambridge, 1981.

[7] A. Rosenthal, *Finite uniform generators for ergodic, finite entropy, free actions of amenable groups*, Probab. Th. Rel. Fields **77** (1988), 147–166.

[8] V. S. Varadarajan, *Groups of automorphisms of Borel spaces*, Trans. Amer. Math. Soc. **109** (1963), 191-220.

School of Mathematics, University of East Anglia, Norwich NR4 7TJ, UK

Department of Mathematics, Ohio State University, Columbus OH 43210, USA